

ABSTRACT

The present invention relates to speech recognition systems, particularly speech-to-text systems and software and decoders for the same. The present invention provides a decoder for an automatic speech recognition system for determining one or more candidate text unit concatenations according to a predetermined criterion and which correspond to a speech segment, the decoder comprising: means for receiving a sequence of feature vectors corresponding to the speech segment; means for mapping with different likelihood values the feature vectors to sequences of nodes in a decoding network, every sequence representing a concatenation of text units; means for determining one or more candidate node sequences in the decoding network corresponding to the candidate text unit concatenations by implementing a dynamic programming token passing algorithm in which each token corresponds to a node and is associated with a number of text unit concatenations and likelihood values for these concatenations, and wherein a token associated with a node in the decoding network is derived from the tokens associated with the previous nodes in the network; wherein tokens from different nodes that are to be passed to a common node are combined to generate a new token corresponding to the common node and associated with an identifier for text unit concatenations and likelihood values associated with the previous tokens of said different nodes. This is combined with means for merging a said token having a said identifier, the text unit concatenations of the said previous tokens being associated with said merged token dependent on their corresponding likelihood values.